MOLECULAR ECOLOGY

Molecular Ecology (2014) **23**, 4362–4372 doi: 10.1111/mec.12864

Chromosomal variation segregates within incipient species and correlates with reproductive isolation

GUILLAUME CHARRON,[1] JEAN-BAPTISTE LEDUCQ[1] and CHRISTIAN R. LANDRY
Département de Biologie, Institut de Biologie Intégrative et des Systèmes, PROTEO, Université Laval, Québec, QC G1V 0A6, Canada

Abstract

Reproductive isolation is a critical step in the process of speciation. Among the most important factors driving reproductive isolation are genetic incompatibilities. Whether these incompatibilities are already present before extrinsic factors prevent gene flow between incipient species remains largely unresolved in natural systems. This question is particularly challenging because it requires that we catch speciating populations in the act before they reach the full-fledged species status. We measured the extent of intrinsic postzygotic isolation within and between phenotypically and genetically divergent lineages of the wild yeast *Saccharomyces paradoxus* that have partially overlapping geographical distributions. We find that hybrid viability between lineages progressively decreases with genetic divergence. A large proportion of postzygotic inviability within lineages is associated with chromosomal rearrangements, suggesting that chromosomal differences substantially contribute to the early steps of reproductive isolation within lineages before reaching fixation. Our observations show that polymorphic intrinsic factors may segregate within incipient species before they contribute to their full reproductive isolation and highlight the role of chromosomal rearrangements in speciation. We propose different hypotheses based on adaptation, biogeographical events and life history evolution that could explain these observations.

Keywords: chromosomal rearrangements, genetic incompatibilities, reproductive isolation, *Saccharomyces paradoxus*, speciation

Received 17 April 2014; revision received 11 July 2014; accepted 14 July 2014

Introduction

Many fundamental questions regarding the molecular mechanisms of speciation and reproductive isolation (RI) remain unanswered (Wolf *et al.* 2010; Butlin *et al.* 2012). Years of research on speciation have revealed that chromosomal changes (CCs) and genetic incompatibilities [negative epistatic genetic interactions, also referred as Dobzhansky–Muller incompatibilities (DMIs)] are important intrinsic mechanisms that cause reproductive barriers (reviewed in Johnson 2010; Presgraves 2010; Wolf *et al.* 2010). Another major question that recently emerged is whether these intrinsic incompatibilities are already present before geographical or other extrinsic factors prevent gene flow between emerging species (reviewed in Cutter 2012). This question is particularly challenging because it requires that we catch speciating populations in the act, that is, identifying species at early or intermediate stages of divergence before they reach the full-fledged species status (Via 2009). Observations in plants and animals (reviewed in Cutter 2012) supporting the presence of polymorphic DMIs within species challenge the assumption that reproductive incompatibilities require the emergence of new incompatible alleles after isolation (Seward 1910; Dobzhansky 1937; Muller 1942). For instance, Corbett-Detig *et al.* (2013) showed that such DMIs are polymorphic within populations of the model species *Drosophila melanogaster*.

There is strong evidence for the role of DMIs and CCs in creating postzygotic incompatibilities. How many and

Correspondence: Jean-Baptiste Leducq and Christian R. Landry, Fax: 1-418-656-7176; E-mails: jean-baptiste.leducq.1@ulaval.ca and Christian.Landry@bio.ulaval.ca
[1]Equal contribution

what type of genes underlie DMIs are major questions that have been addressed in model species (e.g. mice in Mihola *et al.* 2009; *Drosophila* in Phadnis & Orr 2009). The role of CCs in initiating and maintaining reproductive barriers has been noticed early in the study of speciation cytogenetics (White 1978) and they play important roles in particular taxonomic groups such as plants (reviewed in Hoffmann & Rieseberg 2008; Widmer *et al.* 2009). CCs may contribute to species incompatibilities in several ways (Rieseberg 2001; Hoffmann & Rieseberg 2008). First, heterokaryotypes from crosses between individuals with chromosomal differences are often infertile due to the accumulation of chromosomal abnormalities during meiosis, which leads to hybrid sterility (Sturtevant & Beadle 1936; Davisson & Akeson 1993). Another role of CCs is that they enhance the emergence and maintenance of some genetic elements by suppressing recombination (Navarro & Barton 2003; Feder & Nosil 2009), by reducing gene flow between ecologically similar populations (Navarro & Barton 2003; Noor & Feder 2006; Widmer *et al.* 2009) or by enhancing adaptation to new environments (Kirkpatrick & Barton 2006; Hoffmann & Rieseberg 2008; Widmer *et al.* 2009; Lowry & Willis 2010). Finally, large-scale changes such as changes in ploidy may directly contribute to speciation by creating individuals that become reproductively isolated from their parental species within a single generation (Otto & Whitton 2000). This mechanism appears to be important in plants where a large proportion of speciation events are accompanied by ploidy increase (Wood *et al.* 2009).

The role of CCs in speciation may also be important in fungi where there are many reports of extensive CCs among closely related species as well as extensive variation in RI within and among species (Giraud & Gourbiere 2012). Interestingly, many of these CCs may be adaptive, as shown in natural and experimental yeast populations (Gordon *et al.* 2009; Dunn *et al.* 2013). These observations suggest that natural selection could accelerate the fixation of CCs in natural populations and thus contribute to the onset of RI and speciation at the same time. There is strong experimental evidence for a direct causal link between CCs and RI in the model fungal species of the genus *Saccharomyces* and *Schizosaccharomyces*. For instance, Delneri *et al.* (2003) showed that a single chromosomal inversion was sufficient to affect progeny survival in a cross between *Saccharomyces cerevisiae* and *S. mikatae*. In other studies of species of the same genus, both CCs and DMI have been suggested to contribute to RI (Liti *et al.* 2006). More recently, Hou *et al.* (2014) showed that reciprocal translocations play a role in creating RI among strains within *S. cerevisiae*. Finally, Zanders *et al.* (2014) showed that chromosomal rearrangements were playing a role in the RI between *Schizosaccharomyces pombe* and *S. kambucha*, which can mate to form F1 hybrids but produce very few viable F2s. These observations stress the potentially fundamental role of polymorphic CCs in initiating RI among incipient species and show that fungi are powerful and tractable model systems for investigating these effects. However, due to the lack of knowledge on the ecology and biogeography of these species, it is difficult to link the mechanisms of reproductive isolation with the ecological and demographic factors that initiated the process in these organisms (Landry *et al.* 2006; Replansky *et al.* 2008). Moreover, Zeyl (2014) recently pointed out that our current knowledge on the genetic bases of early speciation is mostly limited to few laboratory and mostly inbreeding organisms, stressing the need to investigate these processes in a completely natural context.

We recently identified a new system within natural populations of *Saccharomyces paradoxus*, the sister species of the laboratory model *S. cerevisiae*, that offers an unprecedented case study in ecological and speciation genomics (Fig. 1a). North American *S. paradoxus* populations are composed of three principal lineages: two American lineages (B and C) and a European lineage (A) that was introduced recently in a transoceanic migratory event from a European source population (Kuehne *et al.* 2007; Leducq *et al.* 2014; Fig. 1b). We recently showed that the two American lineages occupy different but overlapping regions (Fig. 1c), are genetically distinct and are phenotypically divergent in a manner that suggests the influence of historical and ecological factors in driving their divergence (Leducq *et al.* 2014). Based on these observations, we propose that the two American lineages of *S. paradoxus* B and C are in an early speciation process. Here, we tested this hypothesis by measuring postzygotic RI between 25 *S. paradoxus* strains representative of the American lineages. Intrinsic postzygotic RI was measured as the fraction of viable progeny within a cross. Genetic divergence globally explained variation in RI measured among strains. In a second step, we performed chromosomal profiling analysis of the parental strains to examine the potential role of CCs in RI. Surprisingly, we found more variation in karyotypic profiles within than among lineages, and this variation correlated with RI measured within lineages. We discuss the potential role of CCs in initiating RI within incipient species and the putative ecological and biogeographical origins of these rearrangements.

Material and methods

Strain collection

We used 25 natural strains of *S. paradoxus* representative of the geographical distribution and of the three

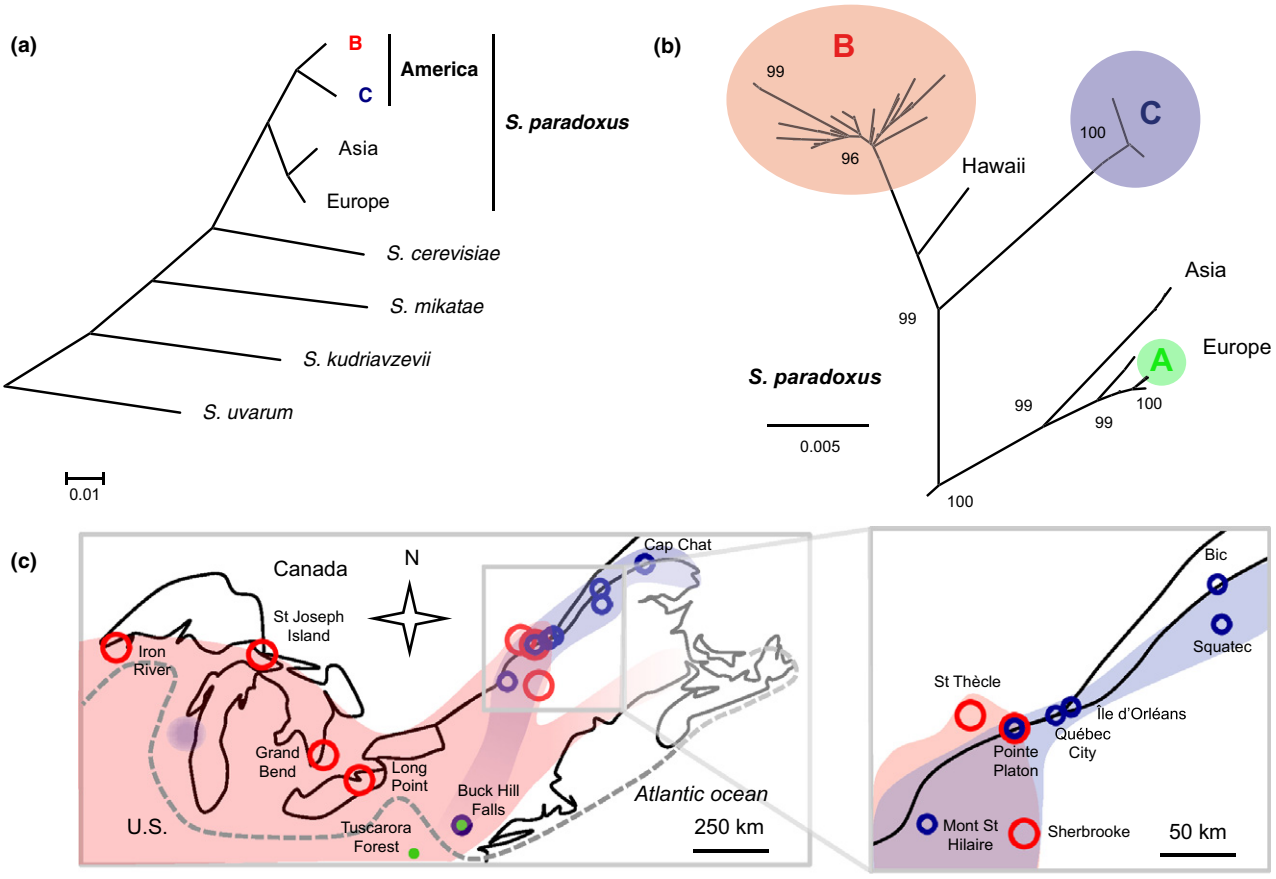

Fig. 1 Two diverging lineages of *Saccharomyces paradoxus* are partially sympatric in North America. (a) Phylogenetic position of main *S. paradoxus* lineages in *Saccharomyces sensus stricto*. (b) Details of the *S. paradoxus* phylogeny. American lineages are denoted as A (recent European introduction; green), B (red) and C (dark blue). The phylogenetic trees are adapted from Leducq *et al.* (2014) (pairwise nucleotide divergence). (c) The geographical distribution of lineages B (red) and C (dark blue) in Northeast America could reflect colonization and secondary contact after the last glaciation event approximately 110 000 to 12 000 years ago (southern limit of the ice *shelf* indicated by a dotted line; Andersen & Borns 1997). Lineage distributions were updated based on additional sampling efforts in 2013 (personal observations; CT Hittinger, personal communication). Gradients indicate that the eastern and southern limits of lineages B and C are defined by rare isolates. The sampling locations of 25 strains used in this study are indicated by coloured circles.

genetic lineages found in Northeast America (Leducq *et al.* 2014): A (4), B (7) and C (14) (Fig. 1c, Table S1, Supporting information). Most strains were isolated on the bark of deciduous trees, and all were isolated according to the same standard protocol (Kuehne *et al.* 2007; Leducq *et al.* 2014).

Heterothallic strain construction and spore viability

We deleted the *HO* gene in wild diploid strains (Table S1, Supporting information) to prevent mating-type switching after sporulation. We amplified *HO* deletion cassettes comprising an antibiotic marker by PCR using oligonucleotides specific to each lineage. We designed lineage A- and B-specific oligonucleotides from available genomic data (Cubillos *et al.* 2009; Liti *et al.* 2009) and lineage C-specific oligonucleotides from the *HO* locus sequence that we obtained from five representative strains (LL2012-006, LL2011-005, LL2011-001, LL2012-016 and YPS667; Leducq *et al.* 2014) after genomic DNA extraction (Amberg *et al.* 2005), standard PCR amplification and Sanger sequencing (NCBI accession nos. KJ410184-88). Cassette amplification, transformation by homologous recombination and cassette integration verification followed Leducq *et al.* (2012), with heat shock performed at 37 °C instead of 42 °C for lineages A and C (oligonucleotides and plasmids are listed in Table S2, Supporting information).

We obtained heterothallic haploid strains by sporulation (Ramaswamy *et al.* 1998) and dissection of the parental diploid strains constructed above. Sporulated cells were dissected (six tetrads per strain) on a YPD (yeast extract, peptone, dextrose) agar plate (2%) using a Sporeplay dissection microscope (Singer, UK) to obtain one spore of each mating type. Mating types were verified by crosses with tester strains of both

mating types with antibiotic resistance markers other than the one of the tested strain. Mating types were inferred from the growth of the strains on YPD agar plate containing both antibiotics.

We performed 87 crosses by optimizing possible pairwise comparisons in terms of geographical distribution and genetic divergence (Leducq *et al.* 2014) among 25 haploid *S. paradoxus* American strains (Table S3, Supporting information). As controls, we included six crosses of strains from each of the three lineages with two other members of the *Saccharomyces* genus, *S. cerevisiae* strain BY4741 (Baker Brachmann *et al.* 1998) and *S. uvarum* strain JRY8147 (Gallagher *et al.* 2009; Table S3, Supporting information). We inoculated 5 mL of fresh YPD medium with 5 μL of each haploid strain to be crossed and incubated them at 30 °C for 24 h with agitation. We selected diploids on YPD agar plates with the two appropriate antibiotics (G418: 200 μg/mL; CloNAT: 100 μg/mL; hygromycin: 250 μg/mL) that were incubated at 25 °C for 48 h and isolated one diploid clone for each cross. We sporulated the clones and dissected 24 tetrads per cross (sporulation and dissection as described above), for a total of 96 possible spores. Dissected spores were incubated at 25 °C for 72 h, and we measured spore survival (*S*) for each cross as the proportion of spores forming a colony visible to the unassisted eye (Fig. 2a).

Karyotype analysis

We prepared yeast chromosomes following Maringele & Lydall (2006) with the following modifications: incubation steps in EDTA-Tris-β-mercaptoethanol and proteinase solutions were conducted overnight, and voltage gradient used for the migration was set to 5.9V/cm. All migrations were performed in a CHEF DR® III variable angle system (Bio-Rad, USA), and the buffer was systematically changed every two migrations. Each gel contained nine samples of *S. paradoxus* (unmodified parental strains) and a *S. cerevisiae* CHEF DNA size standard (Bio-Rad, USA). Gels were stained with ethidium bromide and imaged using a BioRad GEL DOC XR+ molecular imager with the image lab software (Bio-Rad, USA; Fig. 2b). Images were analysed using BIONUMERICS version 7.1 by Applied Math NV (available from http://www.applied-maths.com). Dice's similarity coefficient was used to establish the similarity distance of band patterns based on their position after gel migration. The migration distance of a chromosome is inversely proportional to its length, and a change in chromosomes length by large insertions and deletions translates into a shift in banding pattern. Parameters were determined as to obtain an optimal clustering of the *S. cerevisiae* standard markers. Standard settings were used with the following modifications: optimization 0.5%, tolerance 1.2% and uncertain bands included.

Statistical analyses

Our aim was to test whether we could explain variation in spore survival (*S*) using two potential intrinsic factors of postzygotic reproductive isolation: genetic divergence (*Dg*) and karyotypic divergence based on banding patterns (*Dc*) between parental strains. We estimated *Dg* from two unlinked markers (1460 bp; Leducq *et al.* 2014) using pairwise nucleotide divergence among strains (Maximum Composite Likelihood model; Tamura *et al.* 2011). We used a linear model (*lm*) to explain variation in *S* by considering *Dg* and *Dc* as additive and interacting factors (model M1: *S~Dg+Dc+Dg*Dc*) and estimated the significance of the correlation (Pearson's test) between *S* and each intrinsic factor. We repeated this test by discarding the possible interaction between intrinsic factors, that is, by calculating the correlations between *Dg* and *S* corrected by *Dc* (residuals from *lm*: *S~Dc*) and between *Dc* and *S* corrected by *Dg* (residuals from *lm*: *S~Dg*). To distinguish small- from large-scale effects of intrinsic factors, we repeated the analysis within lineages only (i.e. by excluding interlineage and interspecies crosses; model M2) and among lineages or species only (i.e. by excluding intralineage crosses; model M3). Details of the models are shown in Table S4 (Supporting information). We used Welch *t*-tests to compare averaged *S* among categories of crosses according to genetic lineages. For each test (Pearson's tests, linear models and Welch *t*-tests), we re-estimated the statistics after 100 000 randomizations of *S* values (intrinsic factor fixed to correct for the fact that not all strains were involved in the same number of crosses). We reported *P*, the proportion of randomly re-estimated statistics that were equal or lower than the observed value. Figure S1 (Supporting information) summarizes the method of randomization used, considering that pairwise comparisons between strains were not independent (i.e. strains were involved in different crosses) and unbalanced (i.e. not all combinations of crosses were tested, and some strains were overrepresented). We then considered that the significance of the test was not due to the structure of our data when the observed statistic was significantly higher ($P \geq 0.95$) or lower ($P \leq 0.05$) than the re-estimated statistics. All analyses were conducted in R 3.0.0 (R Development Core Team 2010). All datasets and scripts used are available as Appendices S1 and S2 (Supporting information), respectively.

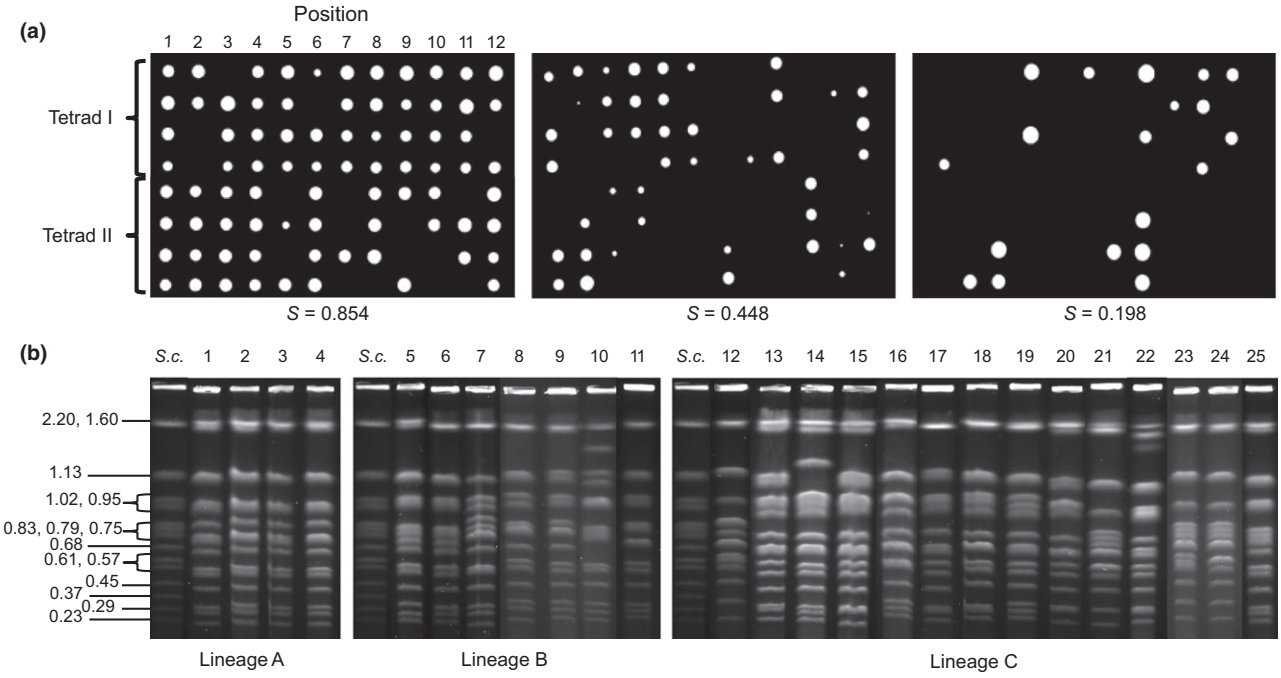

Fig. 2 Reproductive isolation and karyotyping in *Saccharomyces paradoxus* North American strains. (a) Three examples of spore viability (*S* value indicated below) measured as colony growth after tetrad dissection. Each position represents two dissected tetrads (I and II), with four spores for each tetrad (aligned vertically). (b) Karyotypes of strains used in this study as measured by CHEF-PFGE. Each band represents at least one chromosome. Numbers above the karyotypes refer to strains, grouped by lineages (see Table S1, Supporting information). *S.c.* represents *S. cerevisiae* chromosomes standard (chromosome lengths are given in Mbp on the left).

Results

Intra- and interlineage variation in reproductive isolation

We tested whether the North American lineages of *S. paradoxus*, A, B and C (Fig. 1), were postzygotically reproductively isolated. We performed pairwise crosses between strains of lineages A (European lineage but isolated in North America), B and C. Crosses with *S. cerevisiae* and *S. uvarum* yielded low hybrid viability (0–2% of progeny survival; Fig. 3a; Table S3, Supporting information) as reported before (Naumov 1987; Greig 2009). Within *S. paradoxus*, we observed the lowest average spore survival rate (*S*) in crosses between lineages, with 42.1% (A×B), 31.9% (A×C) and 39.7% (B×C) spore viability (Fig. 3a), which is in the range previously observed among lineages (Sniegowski *et al.* 2002; Liti *et al.* 2006; Kuehne *et al.* 2007). B×C crosses showed a significant decrease in average *S* as compared to intra-lineage crosses B×B (58.6%; *P*-value = 0.014, Welch *t*-test) and C×C (61.6%; *P*-value < 0.001, Welch *t*-test; Fig. 3a), and these reductions are significantly stronger than expected over 100 000 random permutations ($P = 0.997$ and $P = 0.999$, respectively; Fig. S2, Supporting information). Surprisingly, we found extensive intralineage reproductive isolation in both lineages B and C, with minimum *S* values as low as 35.4% (B×B) and 19.8% (C×C). This is particularly extreme in lineage C, with some C×C spores having less viability than A×B, A×C or B×C hybrids. This effect was mainly explained by four strains from lineage C from two distinct locations (LL2011_005, LL2011_006, LL2012_016 and LL2012_18; Table S3, Supporting information), which were involved in 16 crosses averaging 42.5% spore viability with other lineage C strains, while they yielded significantly increased spore viability when crossed with each other (86.2%, four crosses; *P*-value < 0.001, Welch *t*-test; Table S3, Supporting information). Removing crosses involving these four strains from the analyses (20 crosses) increases the C×C progeny viability to 73.9%, which remains significantly higher than B×C (*P*-value < 0.001, Welch *t*-test; $P = 0.999$).

Extensive within-lineage chromosomal variation and interlineage nucleotide divergence explain variation in reproductive isolation

We examined whether the extent of reproductive isolation among strains (*S*; Fig. 3a) was correlated with the

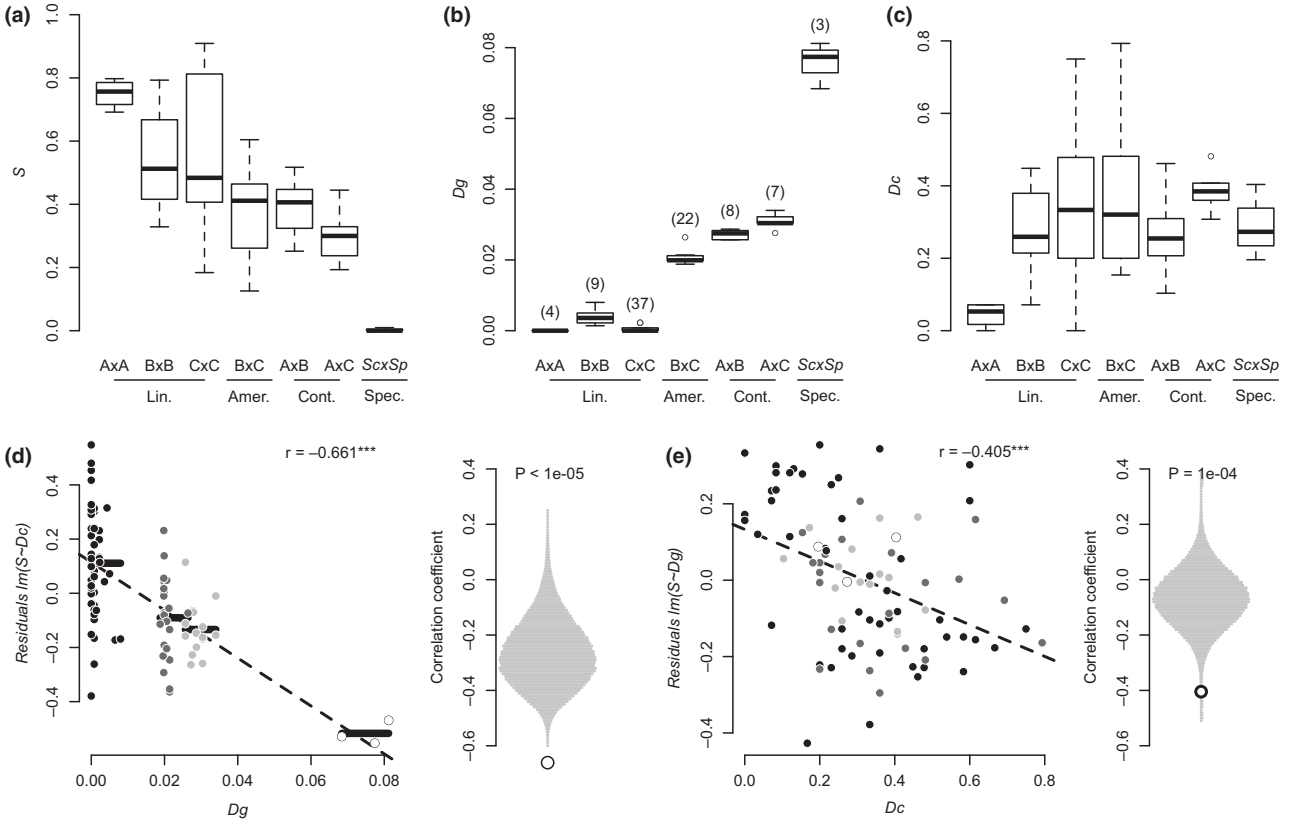

Fig. 3 Postzygotic reproductive isolation (*S*) among 25 *Saccharomyces paradoxus* North American strains correlates with variation among and within lineages in genetic divergence (*Dg*) and chromosome profiles (*Dc*). (a) Boxplot showing the illustrating distribution of *S* among strains within each evolutionary category: within-lineage crosses (Lin.); among American lineages, B×C crosses (Am.); between European and American lineages, A×B and A×C crosses (Cont.); and among species, *S. paradoxus* × *S. cerevisiae* crosses (Spec.). (b) Boxplot illustrating the distribution of *Dg* among strains within each evolutionary category. (c) Boxplot illustrating the distribution of *Dc* among strains within each evolutionary category. (d) *S* significantly decreases with *Dg*. *S* values were corrected by *Dc* (residues from linear model $S \sim Dg$). (e) *S* significantly decreases with *Dc*. *S* values were corrected by *Dg* (residues from linear model $S \sim Dg$). In the left panels of (d) and (e), different colours symbolize different evolutionary categories: crosses within-lineage (black); crosses between American lineages, B×C (dark grey); crosses among European and American lineages, A×B and A×C (grey); and crosses between species, *S. paradoxus* × *S. cerevisiae* (white). Horizontal black bars represent mean *S* value in each category. The dotted line shows the linear regression of *S* as a function of *Dc* (d) or *Dg* (e). Correlation coefficients *r* (Pearson's test) are highly significant ('***': *P*-value < 0.001). Violin plots on the panels indicate that the observed correlation coefficients (black circles) are stronger than expected (grey distribution) over 100 000 random permutations ($P < 0.001$; see Fig. S1, Supporting information).

extent of genetic divergence (*Dg*; Fig. 3b) and differences in their chromosome profiles (*Dc*; Fig. 3c). The extent of variation in *Dc* based on chromosomal banding patterns observed among lineages (0.263–0.387) is similar to what was measured between *S. cerevisiae* and *S. paradoxus* (0.196–0.404; Fig. 3c). Surprisingly, we observed similar amount of variation within lineages B and C (0.044–0.337; Figs 2b and 3c). This is mostly because chromosomal profiles are more variable within lineage C ($Dc = 0.337 \pm 0.207$) than within lineages A and B ($Dc = 0.044 \pm 0.011$ and 0.271 ± 0.089, respectively). This extensive variation in lineage C chromosome profiles was explained by some pairs of strains from lineage C being less similar than some B–C pairs. For instance, the four aforementioned strains (LL2011_005, LL2011_006, LL2012_016 and LL2012_18) have on average more karyotypic similarity with each other ($Dc = 0.080 \pm 0.062$) than with other lineage C strains ($Dc = 0.423 \pm 0.062$; *P*-value < 0.001, Welch *t*-test). We found that globally, *S* was significantly negatively correlated with both nucleotide divergence *Dg* ($r = -0.64$, *P*-value < 0.001, Pearson; Fig. S3a,f, Supporting information) and chromosomal divergence *Dc* ($r = -0.34$, *P*-value < 0.001, Pearson; Fig. S3b,f, Supporting information) among strains. These correlations remain stronger than expected over 100 000 random permutations ($P < 0.001$) and were intensified when correcting *S* by karyotypic divergence (*S*~*Dg*: $r = -0.66$, *P*-value < 0.001, Pearson; $P < 0.001$; Fig. 3d) or by genetic divergence (*S*~*Dc*: $r = -0.40$, *P*-value < 0.001,

Pearson; $P < 0.001$; Fig. 3e). Our linear model (M1) showed that both *Dc* and *Dg* explain a significant part of variation in *S* among crosses (*P*-value < 0.001 in both cases) but that these two parameters are not correlated with each other (*P*-value = 0.130; Fig. S3c, Table S4, Supporting information). The part of variation in *S* explained by *Dg* and *Dc* remains significantly higher than expected over 100 000 random permutations ($P = 0.011$ and $P = 0.004$, respectively; Fig. S4, Supporting information). Analysing separately within-lineage crosses (model M2) from among-lineage/species crosses (model M3) revealed that the intralineage variation in *S* is only explained by chromosomal differences (M2; *Dg*: *P*-value = 0.279, $P = 0.291$; *Dc*: *P*-value < 0.001, $P = 0.005$; Table S4, Supporting information), which was confirmed by a significant correlation between *S* and *Dc* within lineages ($r = -0.46$, *P*-value < 0.001, Pearson; $P = 0.001$; Fig. S3d,f, Supporting information), but not among lineages ($r = -0.061$, *P*-value = 0.707, Pearson; Fig. S3e, Supporting information). The interlineage variation in *S* is only marginally explained by genetic divergence (M3; *Dg*: *P*-value = 0.084; *Dc*: *P*-value = 0.559; Table S4, Supporting information), and the effect is rather due to the structure of data ($P = 0.490$ and $P = 0.513$, respectively; Fig. S4, Supporting information).

Discussion

Intrinsic postzygotic reproductive isolation results from various genetic mechanisms such as differences in ploidy levels, gene–gene incompatibilities, interaction between endosymbionts and the nuclear genome and chromosomal rearrangements (Orr & Turelli 2001; Rieseberg 2001; Coyne & Orr 2004; Noor & Feder 2006; Hoffmann & Rieseberg 2008). From these, gene–gene incompatibilities are a very important mechanism, as they were reported to play a role in hybrid inviability and sterility in both animals and plants (Orr & Turelli 2001; Coyne & Orr 2004). Earlier work on the *Saccharomyces* genus found that nucleotide divergence together with chromosomal changes may play an important role in reproductive isolation via the mismatch repair system (Greig *et al.* 2003). Recent data suggest that cytonuclear interactions may also generate some incompatibilities, resulting in reproductive isolation between more distant species (Chou *et al.* 2010).

Our results show a negative correlation between spore viability and genetic divergence among lineages (Fig. 3d), which supports the accumulation of genetic incompatibilities as a factor in the reproductive isolation among natural lineages of *S. paradoxus* found in North America. This result is consistent with the hypothesis that genetic incompatibilities between lineages accumulate with time and may be correlated with or be caused by global nucleotidic divergence. This result is also in line with the continuum of genetic differentiation and incompatibility that exists between and within species of the *Saccharomyces sensu stricto* group, which may result from the decay in sequence homology along this continuum (Liti *et al.* 2006). The partial intrinsic postzygotic reproductive isolation between lineages B and C may represent an advanced case of this continuum of reproductive isolation. Unexpectedly, we found that a fraction of this reproductive isolation may have been present in the ancestral population of these two lineages and be segregating within lineages or to have accumulated recently, as we observed within-lineage reproductive isolation in both lineages B and C. We observed strains within lineage C that, when crossed with other strains of the lineage, have spore viability in the range of the B×C hybrids. However, crosses among the same strains produced spores with the viability we would expect within a lineage. These strains were isolated in two different locations distant from more than 50 km (Pointe Platon and Île d'Orléans; Fig. 1c, Table S3, Supporting information) and thus represent a part of diversity segregating within lineage C. This suggests that alleles segregating within lineage C contribute to reproductive isolation and that the sorting of these elements is incomplete. Our results support the hypothesis proposed by Corbett-Detig *et al.* (2013) that genetic incompatibilities are common within species.

Chromosome analyses revealed a correlation between reproductive isolation and the divergence of chromosome profiles. Lineage C has more variation in chromosome profiles than the other two lineages, suggesting that it has accumulated more genomic insertions or deletions or rearrangements than other strains. For instance, the four aforementioned strains present about 40% of chromosome divergence with other strains from lineage C, in the same range observed between different lineages or species (Fig. 3c). This correlates with a higher variation in reproductive isolation within this lineage. The evolutionary forces that have contributed to the accumulation of CCs are currently unknown, but several scenarios involving adaptive and nonadaptive forces can be put forward. Genomic rearrangements and CCs in yeasts have been observed as having adaptive roles under stressful environmental conditions. Experimental evolution studies reported that long-time starvation leads to large-scale genomic rearrangements in *S. cerevisiae,* which improved fitness in new starvation episodes (Dunham *et al.* 2002; Coyle & Kroll 2008; Hong & Gresham 2014). In addition, wild *S. cerevisiae* strains isolated from the Evolution Canyon harboured genomic rearrangements that were correlated with their resistance to toxic concentration of copper in their environment (Chang *et al.* 2013). Thus, one possibility is that

subpopulations of the C lineage may have experienced stressful local environmental conditions during the years since they colonized this environment and that these rearrangements were a mean of rapid adaptation (Fig. 4a) and contribute to reproductive isolation among subpopulations as a side effect. Interestingly, some chromosomal rearrangements associated with adaptation to starvation resistance in *S. cerevisiae* were shown to influence the viability of hybrids in experimental evolution settings (Kroll *et al.* 2013), thus coupling the mechanisms of adaptation and reproductive isolation.

Neutral changes may also play a role in the accumulation of CCs. Evolutionary models implying CCs in speciation have been criticized based on the fact that strong genetic drift would be necessary to fix CCs because heterozygous individuals for these CCs are expected to have lower fitness early in the process (Walsh 1982; Lande 1985; Rieseberg 2001). Models developed since then suggested that CCs primarily acted by reducing recombination and that they may have little effects on fitness in heterozygotes (Rieseberg 2001; Navarro & Barton 2003; Noor & Feder 2006). However, the mode of reproduction of budding yeasts may allow for the neutral accumulation of these CCs by genetic drift. The life cycle of the *Saccharomyces* yeasts is comprised of mostly asexual reproduction cycles punctuated by rare sexual reproduction events that, most of the time, end up in the mating of spores from the same tetrad (inbreeding; Tsai *et al.* 2008). The colonization of new environments (Fig. 4a) could have changed these life history traits

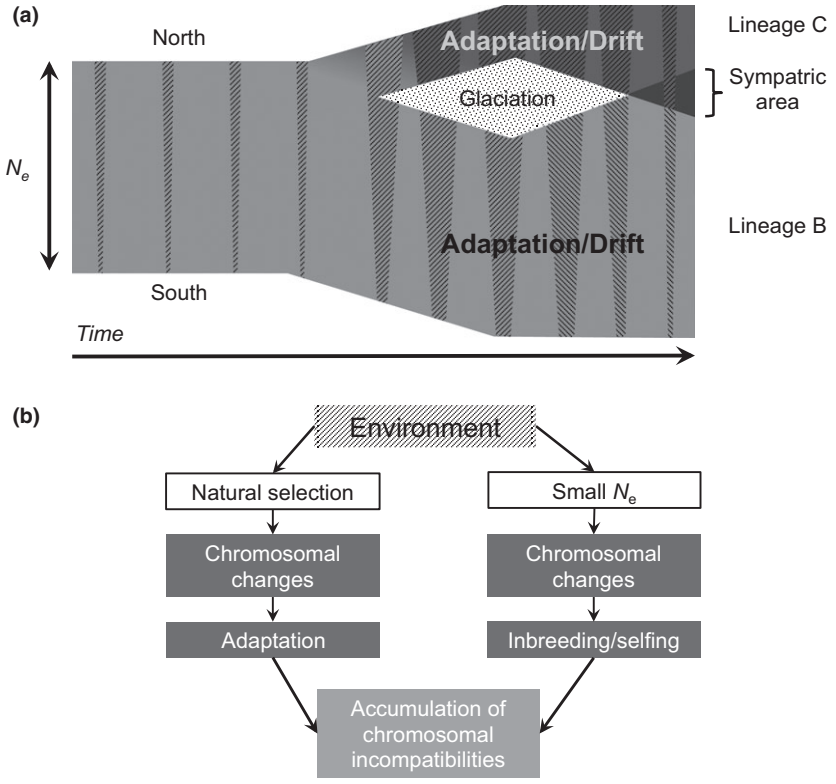

Fig. 4 Proposed scenario for the divergence and secondary contact of the North American lineages of *Saccharomyces paradoxus*. (a) The horizontal timeline represents the evolution of a putative ancestral American *S. paradoxus* population over time (horizontal arrow) into lineages B (grey) and C (dark grey). Lineages may have diverged after allopatric isolation of the lineage C ancestor in a glacial refugium (dotted area). A secondary contact (sympatric area, black) took place after the last glaciation event in the region (110 000–12 000 years ago). Given its northern distribution, lineage C was likely exposed to more frequent and long episodes of extreme conditions (hatched areas) than lineage B. The effective population sizes of the two lineages may have been influenced by these historical and biogeographical events. Dobzhansky–Muller incompatibilities and chromosomal changes may have accumulated through natural selection or genetic drift in each lineage. Selection could have been divergent selection or similar selective pressures that would have fixed alternative alleles. (b) Neutral and selective processes may have led to the accumulation of chromosomal incompatibilities within and between *S. paradoxus* American lineages. A stressful environment at the northern limit of the two lineages may have favoured evolution of adaptive chromosomal changes in local populations. For instance, chromosomal changes could be adaptive through changes in gene dosage (Chang *et al.* 2013) or by the capture of locally adapted alleles (Kirkpatrick & Barton 2006). Alternatively, these environments may have favoured the maintenance of small populations, which favoured inbreeding and made chromosomal changes effectively neutral due to the absence of outcrossing and thus of the formation of heterozygous individuals.

(Nosil 2012) and thus played a role in the accumulation of CCs. Strains of lineage C could mate or outbreed at lower frequencies. Clonal reproduction or sexual reproduction mostly limited to inbreeding would render reproductive incompatibilities among genotypes effectively neutral because heterozygous individuals are rarely produced, leading to their accumulation (Fig. 4b). Two lines of evidence support this model. Lineage C appears to have a more limited geographical distribution than lineage B and a lower density, as suggested by the lower sampling success for *S. paradoxus* strains in the north of its geographical distribution (Charron *et al.* 2013). In addition, average nucleotide diversity in lineage C (0.04+0.01%) is much lower than in lineage B (0.33+0.09%; Leducq *et al.* 2014), consistent with a reduced effective population size. Another possibility is that these changes are generally deleterious and accumulate in lineage C through random genetic drift, without the need for changes in life history.

Both the adaptive (natural selection favouring the accumulation of CCs) and nonadaptive mechanisms (drift allowing the accumulation of CCs) of our models are tightly linked with the recent changes in the habitats occupied by *S. paradoxus* in North America (Fig. 4b). The repeated glaciation events during the Pleistocene have been shown to be a factor that has promoted allopatric speciation via the Mississippian and Atlantic glacial refugia in several taxonomic groups, for instance in freshwater fishes species (April *et al.* 2013), and in *Drosophila* (Ford & Aquadro 1996). This may have also been the case for yeasts, as the break in the B–C distribution corresponds to the biogeographical pattern observed in Northeast American taxonomic groups above (Fig. 1c). Freshwater fishes and some *Drosophila* species are separated in west and east lineages with a secondary contact zone situated along Lake Ontario and the St. Lawrence River, exactly where a geographical co-occurrence zone between B and C was found (Leducq *et al.* 2014). The northern distribution and the putatively low population size of lineage C suggest a colonization event from a glacial refugium, which is also congruent with the evidence of such refugium in the North American ice shelf for deciduous trees on which yeasts reside, for instance *Quercus* (Jackson *et al.* 2000). Historical biogeographical barriers may thus have contributed not only to shape the plant and animal biodiversity of this region but also to shape the microbial one.

Overall, our work suggests that the two North American endemic lineages of *S. paradoxus* are at the initial stage of speciation and thus offer a new model for the study of this process at the molecular level in a genetically tractable and naturally evolving species. The large remaining, yet unexplained, variance in reproductive isolation we observed within lineages also suggests complex intrinsic mechanisms such as polymorphic incompatibilities segregating within lineages. In addition, the mechanisms by which RI could be caused by CCs need further investigation and will require additional analyses that allows finer resolution than the method used here, which can only detect changes that affect chromosome sizes and could miss inversions, which are known to contribute to reproductive isolation in other taxonomic groups (Noor *et al.* 2001; Kirkpatrick & Barton 2006; Hoffmann & Rieseberg 2008; Feder & Nosil 2009; Lowry & Willis 2010). Combined with the genetics and genomics resources of the budding yeasts, this model offers an unprecedented opportunity for the study of the role of evolutionary forces such as natural selection and genetic drift on genes and genome architecture at the early stage of speciation.

Acknowledgements

We thank Emilie Bernatchez, Marie Filteau, Anne-Marie Dion-Côté, Nadia Aubin-Horth, Guillaume Diss and Alexandre K. Dubé for comments on the manuscript. We thank Roger C. Levesque and Iréna Kukavica-Ibrulj for the use of their CHEF-PFGE apparatus. We thank Raghav Nuwal for help in performing crosses during the review process. This work was funded by a NSERC discovery grant to CRL and partially by a HFSP grant to CRL. JBL was supported by a FRQS fellowship and GC by a PROTEO fellowship. CRL is a CIHR's new investigator.

References

Amberg DC, Burke DJ, Strathern JN (2005) Methods in Yeast Genetics: A Cold Spring Harbor Laboratory Course Manual, 2005 Edition (Cold Spring). 119–120.

Andersen BG, Borns HW (1997) *The Ice Age World: An Introduction to Quaternary History and Research with Emphasis on North America and Northern Europe During the Last 2.5 Million Years* Scandinavian University Press North America.

April J, Hanner RH, Dion-Cote AM, Bernatchez L (2013) Glacial cycles as an allopatric speciation pump in north-eastern American freshwater fishes. *Molecular Ecology*, **22**, 409–422.

Baker Brachmann C, Davies A, Cost GJ *et al.* (1998) Designer deletion strains derived from *Saccharomyces cerevisiae* S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. *Yeast*, **14**, 115–132.

Butlin R, Debelle A Kerth C *et al.* (2012) What do we need to know about speciation? *Trends in Ecology & Evolution*, **27**, 27–39.

Chang SL, Lai HY, Tung SY, Leu JY (2013) Dynamic large-scale chromosomal rearrangements fuel rapid adaptation in yeast populations. *PLoS Genetics*, **9**, e1003232.

Charron G, Leducq JB, Bertin C, Dube AK, Landry CR (2013) Exploring the northern limit of the distribution of *Saccharomyces cerevisiae* and *Saccharomyces paradoxus* in North America. *FEMS Yeast Research*, **14**, 281–288.

Chou JY, Hung YS, Lin KH, Lee HY, Leu JY (2010) Multiple molecular mechanisms cause reproductive isolation between three yeast species. *PLoS Biology*, **8**, e1000432.

Corbett-Detig RB, Zhou J, Clark AG, Hartl DL, Ayroles JF (2013) Genetic incompatibilities are widespread within species. *Nature*, **504**, 135–137.

Coyle S, Kroll E (2008) Starvation induces genomic rearrangements and starvation-resilient phenotypes in yeast. *Molecular Biology and Evolution*, **25**, 310–318.

Coyne JA, Orr HA (2004) *Speciation* Sinauer Associates, Incorporated Publishers.

Cubillos FA, Louis EJ, Liti G (2009) Generation of a large set of genetically tractable haploid and diploid *Saccharomyces* strains. *FEMS Yeast Research*, **9**, 1217–1225.

Cutter AD (2012) The polymorphic prelude to Bateson-Dobzhansky-Muller incompatibilities. *Trends in Ecology & Evolution*, **27**, 209–218.

Davisson MT, Akeson EC (1993) Recombination suppression by heterozygous Robertsonian chromosomes in the mouse. *Genetics*, **133**, 649–667.

Delneri D, Colson I, Grammenoudi S *et al.* (2003) Engineering evolution to study speciation in yeasts. *Nature*, **422**, 68–72.

Dobzhansky TG (1937) *Genetics and the Origin of Species*. Columbia University Press, New York.

Dunham MJ, Badrane H, Ferea T *et al.* (2002) Characteristic genome rearrangements in experimental evolution of *Saccharomyces cerevisiae*. *Proceedings of the National Academy of Sciences of the United States of America*, **99**, 16144–16149.

Dunn B, Paulish T, Stanbery A *et al.* (2013) Recurrent rearrangement during adaptive evolution in an interspecific yeast hybrid suggests a model for rapid introgression. *PLoS Genetics*, **9**, e1003366.

Feder JL, Nosil P (2009) Chromosomal inversions and species differences: when are genes affecting adaptive divergence and reproductive isolation expected to reside within inversions? *Evolution*, **63**, 3061–3075.

Ford MJ, Aquadro CF (1996) Selection on X-linked genes during speciation in the *Drosophila athabasca* complex. *Genetics*, **144**, 689–703.

Gallagher JE, Babiarz JE, Teytelman L, Wolfe KH, Rine J (2009) Elaboration, diversification and regulation of the Sir1 family of silencing proteins in *Saccharomyces*. *Genetics*, **181**, 1477–1491.

Giraud T, Gourbiere S (2012) The tempo and modes of evolution of reproductive isolation in fungi. *Heredity (Edinb)*, **109**, 204–214.

Gordon JL, Byrne KP, Wolfe KH (2009) Additions, losses, and rearrangements on the evolutionary route from a reconstructed ancestor to the modern *Saccharomyces cerevisiae* genome. *PLoS Genetics*, **5**, e1000485.

Greig D (2009) Reproductive isolation in *Saccharomyces*. *Heredity (Edinb)*, **102**, 39–44.

Greig D, Travisano M, Louis EJ, Borts RH (2003) A role for the mismatch repair system during incipient speciation in *Saccharomyces*. *Journal of Evolutionary Biology*, **16**, 429–437.

Hoffmann AA, Rieseberg LH (2008) Revisiting the impact of inversions in evolution: from population genetic markers to drivers of adaptive shifts and speciation? *Annual Review of Ecology Evolution and Systematics*, **39**, 21–42.

Hong J, Gresham D (2014) Molecular specificity, convergence and constraint shape adaptive evolution in nutrient-poor environments. *PLoS Genetics*, **10**, e1004041.

Hou J, Friedrich A, de Montigny J, Schacherer J (2014) Chromosomal rearrangements as a major mechanism in the onset of reproductive isolation in *Saccharomyces cerevisiae*. *Current Biology*, **24**, 1153–1159.

Jackson ST, Webb RS, Anderson KH *et al.* (2000) Vegetation and environment in Eastern North America during the last glacial maximum. *Quaternary Science Reviews*, **19**, 489–508.

Johnson NA (2010) Hybrid incompatibility genes: remnants of a genomic battlefield? *Trends in Genetics*, **26**, 317–325.

Kirkpatrick M, Barton N (2006) Chromosome inversions, local adaptation and speciation. *Genetics*, **173**, 419–434.

Kroll E, Coyle S, Dunn B *et al.* (2013) Starvation-associated genome restructuring can lead to reproductive isolation in yeast. *PLoS ONE*, **8**, e66414.

Kuehne HA, Murphy HA, Francis CA, Sniegowski PD (2007) Allopatric divergence, secondary contact, and genetic isolation in wild yeast populations. *Current Biology*, **17**, 407–411.

Lande R (1985) The fixation of chromosomal rearrangements in a subdivided population with local extinction and colonization. *Heredity*, **54**, 323–332.

Landry CR, Townsend JP, Hartl DL, Cavalieri D (2006) Ecological and evolutionary genomics of *Saccharomyces cerevisiae*. *Molecular Ecology*, **15**, 575–591.

Leducq JB, Charron G, Diss G *et al.* (2012) Evidence for the robustness of protein complexes to inter-species hybridization. *PLoS Genetics*, **8**, e1003161.

Leducq JB, Charron G, Samani P *et al.* (2014) Local climatic adaptation in a widespread microorganism. *Proceedings of Biological Sciences*, **281**, 20132472.

Liti G, Barton DB, Louis EJ (2006) Sequence diversity, reproductive isolation and species concepts in *Saccharomyces*. *Genetics*, **174**, 839–850.

Liti G, Carter DM, Moses AM *et al.* (2009) Population genomics of domestic and wild yeasts. *Nature*, **458**, 337–341.

Lowry DB, Willis JH (2010) A widespread chromosomal inversion polymorphism contributes to a major life-history transition, local adaptation, and reproductive isolation. *PLoS Biology*, **8**, e1000500.

Maringele L, Lydall D (2006) Pulsed-field gel electrophoresis of budding yeast chromosomes. *Methods in Molecular Biology*, **313**, 65–73.

Mihola O, Trachtulec Z, Vlcek C, Schimenti JC, Forejt J (2009) A mouse speciation gene encodes a meiotic histone H3 methyltransferase. *Science*, **323**, 373–375.

Muller HJ (1942) Isolating mechanisms, evolution and temperature. In: *Temperature, Evolution, Development* (ed. Dobzhansky T), pp. 71–125. Jaques Cattell Press, Lancaster, Pennsylvania.

Naumov GI (1987) Genetic basis for classification and identification of the ascomycetous yeast. *Studies in Mycology*, **30**, 469–475.

Navarro A, Barton NH (2003) Accumulating postzygotic isolation genes in parapatry: a new twist on chromosomal speciation. *Evolution*, **57**, 447–459.

Noor MA, Feder JL (2006) Speciation genetics: evolving approaches. *Nature Reviews Genetics*, **7**, 851–861.

Noor MA, Grams KL, Bertucci LA, Reiland J (2001) Chromosomal inversions and the reproductive isolation of species. *Proceedings of the National Academy of Sciences of the United States of America*, **98**, 12084–12088.

Nosil P (2012) *Ecological Speciation*. Oxford University Press, Oxford; New York.
Orr HA, Turelli M (2001) The evolution of postzygotic isolation: accumulating Dobzhansky-Muller incompatibilities. *Evolution*, **55**, 1085–1094.
Otto SP, Whitton J (2000) Polyploid incidence and evolution. *Annual Review of Genetics*, **34**, 401–437.
Phadnis N, Orr HA (2009) A single gene causes both male sterility and segregation distortion in *Drosophila* hybrids. *Science*, **323**, 376–379.
Presgraves DC (2010) The molecular evolutionary basis of species formation. *Nature Reviews Genetics*, **11**, 175–180.
R Development Core Team (2010) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
Ramaswamy NT, Li L, Khalil M, Cannon JF (1998) Regulation of yeast glycogen metabolism and sporulation by Glc7p protein phosphatase. *Genetics*, **149**, 57–72.
Replansky T, Koufopanou V, Greig D, Bell G (2008) *Saccharomyces* sensu stricto as a model system for evolution and ecology. *Trends in Ecology & Evolution*, **23**, 494–501.
Rieseberg LH (2001) Chromosomal rearrangements and speciation. *Trends in Ecology & Evolution*, **16**, 351–358.
Seward AC (1910) Darwin and modern science. *Molecular and General Genetics*, **3**, 109–110.
Sniegowski PD, Dombrowski PG, Fingerman E (2002) *Saccharomyces cerevisiae* and *Saccharomyces paradoxus* coexist in a natural woodland site in North America and display different levels of reproductive isolation from European conspecifics. *FEMS Yeast Research*, **1**, 299–306.
Sturtevant A, Beadle G (1936) The relations of inversions in the X chromosome of *Drosophila melanogaster* to crossing over and disjunction. *Genetics*, **21**, 554.
Tamura K, Peterson D, Peterson N *et al*. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution*, **28**, 2731–2739.
Tsai IJ, Bensasson D, Burt A, Koufopanou V (2008) Population genomics of the wild yeast *Saccharomyces paradoxus*: quantifying the life cycle. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 4957–4962.
Via S (2009) Natural selection in action during speciation. *Proceedings of the National Academy of Sciences of the United States of America*, **106**(Suppl 1), 9939–9946.
Walsh JB (1982) Rate of accumulation of reproductive isolation by chromosome rearrangements. *American Naturalist*, **120**, 510–532.
White M (1978) Chain processes in chromosomal speciation. *Systematic Biology*, **27**, 285–298.
Widmer A, Lexer C, Cozzolino S (2009) Evolution of reproductive isolation in plants. *Heredity (Edinb)*, **102**, 31–38.
Wolf JB, Lindell J, Backstrom N (2010) Speciation genetics: current status and evolving approaches. *Philosophical Transactions of the Royal Society of London. Series B, Biological sciences*, **365**, 1717–1733.
Wood TE, Takebayashi N, Barker MS *et al*. (2009) The frequency of polyploid speciation in vascular plants. *Proceedings of the national Academy of sciences*, **106**, 13875–13879.
Zanders SE, Eickbush MT, Jonathan SY *et al*. (2014) Genome rearrangements and pervasive meiotic drive cause hybrid infertility in fission yeast. *eLife*, **3**, e02630.
Zeyl C (2014) Evolution: a collection of misfits. *Current Biology*, **24**, R394–R396.

C.R.L., J.B.L. and G.C. designed the experiments. G.C. carried out the experiments. J.B.L. and G.C. conducted the analysis. G.C., J.B.L. and C.R.L. wrote the article.

Data accessibility

Raw data and scripts for the statistical analyses of spore viability are provided as supporting information (Appendices S1 and S2).

NCBI accession nos. for *HO* sequence in lineage C: KJ410184-88.

Supporting information

Additional supporting information may be found in the online version of this article.

Fig. S1 Principle of permutation tests used in this study to control for the effects of non-independence of pairwise comparisons and of bias in strain usage in pairwise comparisons.

Fig. S2 Violin plots showing the expected distribution (grey lines) of *t*-values from Welch *t*-tests comparisons of *S* values among categories of crosses (see Fig. 3a).

Fig. S3 Spore survival (*S*) is negatively correlated with genetic divergence (*Dg*) and chromosomal divergence (*Dc*).

Fig. S4 Violin plots showing the expected distribution (grey lines) of estimated coefficients from linear models (see Table S4).

Table S1 List of yeast strains used in this study.

Table S2 List of oligonucleotides and plasmids used for *HO* locus deletion.

Table S3 List of crosses realized in this study.

Table S4 Results from linear model (*lm*) of spore survival (*S*) variation in function of genetic divergence (*Dg*) and karyotypic divergence (*Dc*) and their interaction (*Dg*Dc*) as additional factors (*lm Formula = S ~ dg+dc+ dg*dc*).

Appendix S1 R script used to perform statistical analyses of spore survival, to generate Fig. 3, Fig. S1-S4 and Table S4. Input is provided in Appendix S2.

Appendix S2 Input for statistical analyses of spore survival (R script provided in Appendix S1). Contain raw data of spore survival (S), karyotypic divergence (Dc) and genetic divergence (Dg).